SECOND SEMIANNUAL STATUS REPORT

OF

NASA RESEARCH GRANT NO. NsG-568

<u>Title of Project</u>: Stochastic Models for Multi-valued, Multi-dimensional Relations

<u>Principal Investigator</u>: Triloki N. Bhargava

<u>Period Covered</u>: April 16, 1964, to October 15, 1964

UNPUBLISHED PRELIMINARY DATA

<u>Date</u>: October 31, 1964

Signature T. N. Bhargava
of Principal Investigator

<u>Title</u>: Associate Professor of Mathematics and Statistics

<u>Address</u>: Kent State University
Kent, Ohio 44240

PRELIMINARIES

This is intended to be the second semiannual status report, covering the period from April 16, 1964, to October 15, 1964, of the NASA research grant number NsG-568 for support of the research project entitled "Stochastic Models for Multi-valued, Multi-dimensional Relations." The first semiannual status report, covering the period up to April 15, 1964, has already been submitted earlier.

For the sake of clarity, this report is divided into the following Parts:

I. Chronological Summary

II. Topical Summary

1. Further Progress on Work Reported Previously
2. Introduction and Plans for Research on New Problems

III. Appendices

1. Abstracts
2. Information Theory, Entropy, and Probability Measures
3. Resumês

Abstracts of the four technical reports cited herein are contained in the Appendices; full drafts of these reports are included as separate enclosures.

I. CHRONOLOGICAL SUMMARY

The following represents a short summary of work completed during the report period, divided on the basis of academic quarters.

A. April 16 to June 15, 1964: Research assistants Ahlborn and Curtis continued working while supported by NASA, and graduate assistants Ohm and Wilson maintained an active interest in the NASA project although not on NASA assistantships. Bhargava and Ahlborn presented a joint paper entitled "Directed Graphs and Point-Set Topology" at the regional meetings of the American Methematical Society and the Mathematical Association of America at New York City, April 20-23, 1964. Ahlborn completed work on his master's thesis, entitled "On Directed Graphs and Related Topological Spaces" (Technical Report No. 2). Curtis continued an extensive and exhaustive study of the theory of digraphs, while concurrently workin g on his master's thesis, "On Some Counting Problems in Digraph Theory." Three well-known mathematicians visited during this quarter to discuss problems of mutual interest related to the NASA project. Doyle stayed from April 18 to April 24; Gani, May 7 and 8; and Chatterji, May 22 to May 31.

B. June 16 to September 15, 1964: The principal investigator worked full time on the project during this period, devoting part of the time to organizing the work already done and part of the time to some new problems. Bhargava and Ahlborn started writing a joint paper entitled "Directed

Graphs and Point-Set Topology" (Technical Report No. 1) for publication purposes. Ahlborn, however, was awarded an NSF Summer Fellowship and spent most of his time on that study program. Curtis, due to pressing personal circumstances, decided to engage in full-time teaching, so that his work on counting problems in digraph theory has been temporarily delayed and he has yet to finish his master's thesis. With the exception the first two weeks in August, Wilson made an intensive, full-time study of the problem of developing an information-theoretic notion of entropy for graphs. During this time she was partially supported by NASA funds. Ohm, studying under an NSF Summer Fellowship, began investigating digraphs from the point of view of algebraic structures. Chatterji visited from September 1 to September 21, working with Bhargava on some problems in random graphs and digraphs and on counting topologies. Together they also noticed certain interesting tie-ups between information-theoretic notions and other mathematical theories and work was begun on these problems. They prepared jointly an invited paper, entitled "Random Graphs and Digraphs" (Abstract E) which was presented by Bhargava, at the Joint European Conference of the Institute of Mathematical Statistics and other statistical societies at Berne, Switzerland, in September, 1964.

C. <u>September 16 to October 15</u>: Only one research assistant, Wilson, was appointed to work on the NASA project, with support guaranteed by the Office of the Dean for Research at Kent State University. However, Ohm also continued working on project-related research problems, supported by an NSF Cooperative Graduate Fellowship. Another graduate student, Flagg (for resumé, see III.3) showed interest in working on the project and was assured of an NASA assistantship, provided the grant was renewed. Bhargava continued work on the paper being prepared jointly with Ahlborn and submitted a revised version of his own paper entitled "Time Changes in a Digraph" (Technical Report No. 3) for publication in the Journal of Applied Probability (to appear in June, 1965). Also, Chatterji nearly completed his paper "On Counting Topologies" (Technical Report No. 4), which is to be submitted for publication to The American Mathematical Monthly.

II. TOPICAL SUMMARY

In this section the progress made on the research work originally reported in the first semiannual status report is briefly summarized and descriptions of new problems being studied is given.

II.1 Further Progress on Work Reported Previously.

The following topics were introduced in the first semi-annual status report. Both the new progress made and indications of different approaches and future plans are discussed below:

A. A Link-up Between Directed Graphs and Point-Set Topology: A digraph topology has been definitely established which shows great promise for future study, and a research paper on this interesting relationship between the theory of digraphs and point-set topology is being submitted for publication purposes. The preliminary report was presented at the joint regional meetings of AMS and MAA at New York City in April, 1964, by Ahlborn, a research assistant, jointly with Bhargava, the principal investigator. Ahlborn also wrote his master's thesis in the same general field. Detailed results are to be found in Technical Reports No. 1 and 2. Currently, Bhargava is making further investigations in this direction, with a view to establishing probability structures on the digraph topology.

B. Counting Topclogies: Although it has not yet been possible to solve this rather formidable problem completely, a powerful asymptotic approximation has been obtained using the notion of chromatic graphs. Part of these results (Abstract E) were presented jointly by Bhargava and Chatterji at the Joint European Conference in Berne, Switzerland, and an individual research paper Chatterji has been written for publication purposes (Technical Report No. 4).

C. Most General Way of Describing a Graph and a Digraph: Although no directly significant progress has been made on this problem, various indirect and incidental results have been obtained which should contribute to a continuing investigation of this kind.

D. Random Graphs and Digraphs: Some of the results from the preliminary report have already been presented at the Berne meetings, and more results are still being obtained. Flagg, presently a graduate assistant, is studying a new approach--that of the evolution of a random graph as a stochastic process, in particular, as a birth-and-death process. It is expected that this line of investigation will continue later either under the direct support of NASA or as part of a master's thesis of Flagg. A different approach on the counting problems in random digraphs is described in II.2.I.

E. Information Theory and the Theory of Graphs: Although no new significant results have yet been obtained, preliminary investigations have revealed various possibilities, notably certain interesting tie-ups between information theory and other mathematical theories. These are described more fully in II.2.H and III.2. Also, the NASA research assistant, Wilson, is investigating information-theoretic concepts as more specifically related to the theory of graphs, in an attempt to develop a notion of entropy for graphs, which as shown by Bhargava serve as mathematical models in certain applications.

F. <u>Historical Survey of the Theory of Digraphs</u>: No work has been done on this since Curtis left the project, mainly because other students on the project are working primarily on different problems of particular interest to them. It is hoped that this survey might be completed sometime in the future.

G. <u>Digraphs and Weak Metric</u>: Ahlborn having joined the University of Rochester for his doctoral work, this problem remains virtually in its original state, but the principal investigator, Bhargava, is planning to work in this direction himself, especially because it is thought that the notion of weak metric in digraphs may lead to some interesting information - theoretic results.

II.2. Introduction and Plan for Research on New Problems

In addition to the research continuing on the topics discussed in the previous section, new approaches have developed in relation to certain problems which are explained in more detail in this section:

H. <u>Information Theory, Entropy, and Probability Measures</u>: The research now underway in this area is primarily concerned with an investigation of the different links among various purely mathematical theories through utilization of the information concept and of important applications of information-theoretic notions to the fields of probability, communication theory, and the theory of graphs. Although the concept of information is relatively new in the domain of mathematics, its use and applications have already been shown to be manifold and far-reaching. The main objectives of this research are (i) to develop a suitable notion of entropy in certain kinds of stochastic processes and its relationship to the notion of entropy in ergodic theory, (ii) to explain the occurrence of information-theoretical quantities in connection with the Hausdorff dimension and entropy in probability distributions, and finally (iii)to show applications of the concept of information in statistics and communication theory. A more complete description of the research being conducted in this field (jointly with Chatterji) and a more detailed explanation of some specific problems under investigation will be found in Appendix III.2. Also, a further discussion of information theory as related to directed graphs is given in II.K.

I. Random Graphs and Digraphs: Let n be the number of vertices, and N=f(n) be the number of edges, $o \leq N \leq \binom{n}{2}$. Let $P_{n,N}$ or $P_{n,f(n)}$ be the probability of obtaining a completely connected random graph with n vertices and N=f(n) edges.

There seem to be two different ways of looking at $P_{n,N}$:

(i) Gilbert's approach - For every pair of points (i,j) we define a random variable

$$X_{ij} = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1-p=q \end{cases}$$

where all X_{ij}'s are independent. The number N, or f(n), the number of edges, is also a random variable. Gilbert shows that

$$P_{n,N} \sim 1 - nq^{n-1}$$

(ii) Erdos and Renyi's approach - The number of edges, N, or f(n), is fixed. If $N=[n/2 \log n + cn]$, then they show that

$$P_{n,N} = e^{-e^{-2c}}$$

We suggest that $P_{n,N}$ is a function of n and N and assume that:

(a) $n - 1 \leq f(n)$ and $f(n) \leq \frac{n-1}{2}$

(b) f(n) is monotonically increasing

(c) $\{f(n) + 1) /f(n)\}\to 1$.

Interesting results are then obtained regarding different kinds of bounds for f(n) and $P_{n, f(n)}$.

It is planned to compare further the respective methods of Gilbert and of Erdos and Renyi, with particular regard to their extensions to directed graphs.

J. Digraphs and Algebraic Structures: A digraph may be thought of as a set of points with a relation which associates certain ordered pairs of points with one another by means of directed edges. From this point of view there would seem to be a certain analogy between a digraph and a groupoid, a simple algebraic structure consisting of a set of elements and a binary operation defined on certain ordered pairs of elements. Research has begun in the direction of experimenting with various defining relations between graphs and groupoids and of investigating what identification theorems might be obtained from each. Presently, the following definitions seem most promising: Given that H is a halfgroupoid and G(A) is a digraph, where {a,b,c.....} is the set of elements of H and vertices of G(A), then

(i) allow a directed edge from a to b in G(A) if $a \cdot b = c$ in H

(ii) allow a directed path from a to b in G(A) if $a \cdot b = c$ in H

(iii) allow a directed edge from a to c in G(A) if $a \cdot b = c$ in H.

It is easily seen that (i) determines a biunique mapping between the edges of G(A) and the ordered pairs of elements for which the operation is defined in H, but it is limited by the fact that a groupoid corresponds to a complete graph. This handicap is somewhat overcome by (ii), but this definition leads to problems of non-uniqueness. Finally, although (iii) is a many-to-one mapping of H into G(A), it does suggest several interesting relationships. For example, a cyclic groupoid yields a graph possessing a Hamiltonian line whose length equals the order of the groupoid; prime elements in H correspond to unaccessible points in G(A) and idempotent elements to loops; also, there seems to be some relationship between an antigroupoid and a kernel of a graph. Thus, to date, (iii) appears to have the most potential of the three as a basic definition to correlate concepts of graphs and groupoids. It is planned to continue investigation of all of these definitions and of various others which have been formulated, in the hope of finding one that will establish a really useful correspondence between the two systems. Because of the similar conceptual natures of graphs and groupoids, it is felt that significant results can be obtained along these lines and that this approach has good potential for continuing successful research.

K. <u>Entropy and Information-theoretic Concepts in Theory of Graphs</u>:

The concept of entropy, which has emerged from the field of communication theory, has been approached from two essentially different points of view by its major investigators. The axiomatic approach is characterized by Shannon's postulates for a measure of information. He defines entropy of a probability distribution $P = (p_1, p_2, \ldots. p_{\underline{a}})$ as the amount of information contained in a single observation of a random variable X which takes on $\underline{a}$ different values $x_1, x_2 \ldots.. x_{\underline{a}}$, with probabilities $p_k = P(X=x_k)$, $(k=1,2, \ldots.. \underline{a})$. Shannon's postulates pointed to a logarithmic function, and he defines the entropy, H(P), to be

$$H(P) = \sum_{k=1}^{a} p_k \log (1/p_k)$$

On the other hand, Wolfowitz approaches the problem from the pragmatic standpoint, starting with certain particular problems of information theory and accepting as measures of the amount of information the quantities which present themselves in the solutions. Renyi contends that these are not opposing points of view but rather are compatible and complement each other.

The direction of present investigation is to relate the idea of entropy, most probably as derived by Shannon, to the theory of graphs in order to find some measure of information on a random graph, and in particular, on a random directed graph (digraph).

In considering a digraph of order N as a probabilistic model, one may establish a probability, distribution in relation to the length of a directed path from P_i to P_j. It is known that an element of the matrix C^k (where C is the incidence matrix corresponding to its digraph with ones in the principal diagonal), $k > 0$, corresponds to the existence or otherwise of directed paths of length at most k in the corresponding digraph. A rather general probability scheme for our purposes may be defined in the following manner: For an ordered pair (i,j), let $P(C_{ij}^{(k)} > 0) = pK_{ij}$ be the probability that the (i,j)-th element of the matrix C^k is greater than zero (where $C_{ij}^{(k)} > 0$ implies that there is a directed path of some length connecting P_i to P_j;

$(i,j = 1,2,\ldots N)$, $k=1,2,\ldots,N-1$. Various simplifications of this probability distribution as well as some other probability distributions are being considered in order to investigate some basic concepts of entropy for such a digraph.

At present one of the most promising directions of investigation consists in describing the entropy in terms of some function of the sum of the incidence matrix and its powers, that is,

$$(I - C)^{-1} = I + C + C^2 + \ldots + C^k + \ldots$$

where I is the NxN identity matrix, and where C^k is known to be zero for all $k \geq N$, in case of digraph of order N. A more general approach may consist in attaching a suitable weight, say ω, and its powers (corresponding to the length of the directed chain) to the matrix C and its power, i.e. describing the entropy in terms of some function of $(I - \omega C)^{-1}$.

III. APPENDICES

1. ABSTRACTS
2. INFORMATION THEORY, ENTROPY, AND PROBABILITY MEASURES
3. RESUMÉS

III.1. ABSTRACTS

A. Abstract of Technical Report No. 1: Let $\Gamma(A)$ be a digraph defined on the set A, consisting of N points. Let us first consider $S = [\Gamma(a), a\epsilon A]$, the set of all subdigraphs of $\Gamma(A)$ on n points. We may define partitions on S based on various notions of accessibility and obtain identification and counting theorems, derived in terms of an equivalent representation of $\Gamma(A)$ as an incidence matrix (Abstract 64T-138, AMS Notices, February, 1964). Let us next consider $\varepsilon = [E; E \leq A \times A]$, the set of all digraphs on the set A. We define a set $B \leq A$ to be open with respect to E if for all $i\varepsilon (A - B)$, $j\varepsilon B$, we have $\langle i, j\rangle \varepsilon (A \times A - B)$. This determines a topology $(A, \mathcal{T}_E)$ on $\Gamma(A)$ which has complete additive closure. We note various other interesting properties of this topology, such as: $(A, \mathcal{T}_E)$ is T_1 if and only if $E = \phi$ (Abstract 611-80, AMS Notices, April, 1964). Then some results are obtained in terms of both partitions and the digraph topology by introducing the notion of the "core" of a point - the intersection of the smallest open and closed sets containing the point - and the "core function." Also, certain statistical results for digraphs are obtained by introducing probability measures on S. Finally, a discussion of chromatic graphs is presented, with some of the results for digraphs being extended to the general case.

B. Abstract of Technical Report No. 2: In this report, we investigate certain tie-ups between the theory of directed graphs and point-set topology. This work extends certain aspects of the work done by Bhargava in "A Stochastic Model for Time Changes in a Binary Byadic Relation."

With each directed graph $\Gamma(A,E)$ on an arbitrary set A, we associate a unique topological space (A,T) by defining a set $a \subseteq A$ to be open if there does not exist an edge in $\Gamma(A,E)$ from set $(A - a)$ to set a; each such topology is shown to have the property of completely additive closure. We obtain several theorems relating connectedness and accessibility properties of a directed graph to properties of the topology determined by that directed graph. It is found that the connectedness of a directed graph is, in a certain sense, consistent with the "topological connectedness: of the topological space determined by that directed graph. We further investigate these topologies in terms of the closure, kernel, and core operators.

We show that the definition of an open set, as given above, establishes a single-valued mapping of the family of all topologies with completely additive closure on set A. This mapping also maps one-to-one the family of all transitive directed graph with loops on set A onto the family of all topologies with complete additive closure on set A. Furthermore, the transitive directed graph with loops mapped to a particular topology is, in each case, the directed graph with the maximum edge set determining that topology. On the other hand, we find that there does not necessarily exist a directed graph with a minimal edge set determining a particular topology.

Finally we make a brief study of the properties of topologies obtained from a directed graph with respect to two other definitions for an open set.

C. Abstract of Technical Report No. 3: A probabilistic model for analyzing time changes in the ordered binary dyadic relation on a finite set of points is presented by treating the total relation on the set as the aggregate of its subrelations of fixed order. These subrelations may then be classified in various ways: One system discussed herein is based on the notion of connectivity, and the other on the notion of accessibility. The process of change in the total relation over time is shown to be a time-dependent process taking values on the digraph, or, equivalently, on the incidence matrix isomorphic with the relation. Certain formal procedures for counting these transitions are also set forth, and standard statistical analyses are made. Under suitable assumptions, the probability model presented here has potentialities for application in those situations which can be represented mathematically in terms of a finite set of points and an all-or-none relationship between ordered pairs of these points, Some examples are group dynamics, communication networks, ecology, animal sociology, and management sciences.

D. Abstract of Technical Report No. 4: Let $\Omega_n = [1,2,3...n]$. A topology on Ω_n is a collection T of subsets of Ω_n such that (1)ϕ, $\Omega_n \varepsilon T$ and (2) $A, B, \varepsilon T \Rightarrow A \cap B \varepsilon T, A \cup B \varepsilon T$. Let t(n) be the number of distinct topologies on Ω_n.

We have tried to solve the problem of obtaining an analytic expression for $t(n)$. The problem seems quite hard, and we have only been able to obtain certain inequalities for $t(n)$. We have also noticed a connexion with counting the number of graphs of a certain type.

It can be shown that a topology on Ω_n can be specified equally well by a mapping C from Ω_n to $P(\Omega_n)$ - the power set - or set of subsets of Ω_n. The mapping C must have the following properties:

1. For all $i \epsilon \Omega_n$, $i \epsilon C(i)$
2. If $j \epsilon C(i)$ then $C(j) \subset C(i)$.

$C(i)$ can be thought of as being the closure of the point "i". That such a function can be used to establish a topology on Ω_n may be verified by using the Kuratowski closure postulates. The problem, then, is to count the number of such mappings.

Clearly $C(i)$ can be chosen in 2^{n-1} ways. So $t(n) \leq 2^{n(n-1)}$. On the other hand, the number of algebras of subsets on Ω_n - call it $a(n)$ - is clearly less than, or equal to, $t(n)$. Asymptotic expressions for $a(n)$ can be written down explicitly using number-theoretic partition-functions. Thus we have: Theorem: $a(n) \leq t(n) \leq 2^{n(n-1)}$ (Note: The r.h.s. of the inequality is strict except for n=1 and 2. The l.h.s. is strict for $n>1$).

It is known (Bhargava and Ahlborn) that one can set up a one-to-one correspondence between the topologies on Ω_n and the number of accessibility digraphs on n vertices. (An accessibility digraph is one in which either of two vertices are joined by a direct edge or else there is no string of properly directed edges connecting them).

Many other problems of this type of counting can be posed. For example, how many topologies are T_o (there is only one T_1 topology), how many connected, how many non-homeomorphic, etc. Further, what proportion of topologies in $\Omega_n \times \Omega_n$ are product topologies?

Further results on t(n), the number of topologies on n points, are obtained. For example, it is shown that there exists a number $c > o$ such that for any $\varepsilon > o$,

$$\frac{t(n)}{2^{(c+\varepsilon)n^2}} \to 0, \quad \text{but} \quad \frac{t(n)}{2^{(c-\varepsilon)n^2}} \to \infty$$

Further bounds on c have been obtained. For example, $1/4 \leq c \leq 1/2$. It is conjectured that $c = 1/2$.

The results obtained were derived from a generalization of the idea of chromatic graphs, work on which is proceeding.

E. Abstract; "Random Graphs and Digraphs": We have extended the study of random (undirected) graphs as initiated by Erdös and Rényi (See in particular their paper "On the Evolution of Random Graphs," Publ. Math. Inst. Hungar. Acad. Sci. 5 Ser. A. 1960) to the study of directed graphs (called digraphs). Various models of randomness are possible and we discuss each one, correlating our models with the ones used in the undirected theory (e.g. by Gilbert, Austin, Fagen, Penney, and Riordan). We have concentrated on obtaining only asymptotic results instead of closed form formulae since (as in Erdös and Rényi) the asymptotic results offer a good qualitative picture of the evolution of graphs. (Published in Annals of Mathematical Statistics, Vol. 35, September, 1964).

III.2. Information Theory, Entropy and Probability Measures

The concept of information obtained from the realization of a random experiment was first introduced by Shannon [1] in the year 1948. Although it is a relatively new concept in the domain of mathematics, its use and applications in mathematics have been manifold and far-reaching. We mention in passing the applications in (and indeed the development of) coding theory in the theory of communications, (for example, Shannon and Weaver [2], one of the early works, elaborations and extensions of which are far too many in number to be included in a passing reference). Also noticeable is the impact of the information concept in formulating and reorienting basic ideas in statistics (for example, Kullback [3]). Our interest in information theory consists in studying and exploiting the different link-ups among various purely mathematical theories (with applications) which have been brought to light recently through the usage of the information concept.

To be more specific, we list the different subjects which interest us and which have been vitually touched by information-theoretical notions.

(i) A certain class of probability measures which are naturally induced into the unit interval by a single discrete process (a finite state, space stationary Markov chain) turns out to be singular. The Hausdorff dimensions of the subsets on which these measures 'sit' were calculated by Kinney [5] and Billingsley [6] and, surprisingly, these dimensions turn out to be information on suitable spaces. One of us (see Chatterji [7]) extended the above-mentioned results in certain directions and found the dimensions connected, as observed before, to information. Kinney and Billingsley both try to describe, using coding theory, why these information quantities turn up, but their explanations, we feel, need to be extended to explain fully this remarkable occurrence of information-theoretical quantities in this connexion. We briefly mention in passing that the notion of capacity which is connected to the theory of fractional dimensions is also relevant here and needs to be studied.

To be more specific, let $E \subset [0,1]$, and $\rho > 0$. For $\varepsilon > 0$, let

$$L_\rho(E,\varepsilon) = \inf \{\sum_i |A_i|^\rho : \cup A_i \supset E,\ |A_i| < \varepsilon\}$$

where A_i's are intervals, and $|A|$ denotes the Lebesgue measure of A.

Let $L_\rho(E) = \lim_{\varepsilon \to 0} L_\rho(E,\varepsilon)$. Then $L_\rho(E)$ is the ρ-dimensional Hausdorff outer-measure of E, and the Hausdorff dimension of E equals

$$\inf \{\rho : L_\rho (E) = 0\} \text{ or, equivalently, } \sup \{\rho : L_\rho (E) = \infty\}$$

It can be shown relatively easily (and is relevant for our discussion) that, in the definition above, the covering intervals may be restricted to intervals with binary-rational endpoints without changing the dimension.

From the general nature of the definitions of "entropy of a source" and "capacity of a channel" in the modern theory of information, (see Khinchin [10]) it becomes plausible to conjecture that the Hausdorff dimension of sets on which certain measures sit are connected with the notion of capacity of a channel. For elaboration, see Kinney [5] and Billingsley [6]. Corroborating evidence is to be found in Chatterji [7]. A general theory is greatly desirable. As a beginning, one would study these measures on the unit interval which are induced into it by stationary stochastic processes taking only a finite number of values (use decimal mapping here) or countable infinity of values (use continued fraction mapping here).

(ii) Rényi [8] has introduced a notion of dimension and entropy of probability distributions. Our investigations have shown that in certain instances Rényi's notion of dimension coincides with the classical notion of Hausdorff dimensions. Also to be remarked is the fact that Rényi's notion of entropy leads again to information-theoretical quantities. This area needs to be investigated also.

To be more specific, let μ be a probability measure on $(-\infty,\infty)$, and let

$$H_n(\mu) = \sum_{k=-\infty}^{\infty} \mu(A_{kn}) \log \mu(A_{kn})$$

where $A_{kn} = \{x \mid \frac{k}{n} \leq x < \frac{k+1}{n}\}$

If $\lim_{n\to\infty} \frac{H_n(\mu)}{\log n} = d$ exists, then μ is said to have dimension d.

If d is the dimension of μ, and if

$\lim_{n\to\infty} \{H_n(\mu) - d \log n\} = h$ exists, then "h" is called the entropy of μ. (See Rényi [8]).

Rényi showed that if μ is absolutely continuous and $H_1(\mu) < \infty$ then $d = 1$, and

$$h = \int_{-\infty}^{\infty} f(x) \log \frac{1}{f(x)}\, dx,$$

where f(x) is the density of μ. If μ is purely discrete, $d = 0$ and a similar result holds. Our investigations have shown that for certain continuous singular μ, "d" turns out to be the Hausdorff dimension of a set on which μ is centered. This is perhaps not too surprising in view of known theorems concerning the relation between Lipschitz conditions satisfied by measures and the Hausdorff dimensions of the sets on which they are based. A general theorem connecting "d" and Hausdorff dimensions is desirable but seems difficult to obtain.

(iii) Brgodic theory (see Halmos [4]: A notion of the entropy of a transformation on a probability space has been introduced by Kolmogorov, who was obviously motivated by information theory. This notion helped him in solving a long, outstanding problem in ergodic theory.

(iv) Kolmgorov has introduced a concept of E-entropy in compact metric spaces which is a close analogue of information-theoretical concepts (see Lorentz [9] for an exposition). We feel that investigations to convert this analogy to a logical unity will be fruitful.

It is hoped that the proposed research will provide an explanation of the occurrence of information-theoretical quantities in various branches of pure mathematics. This, we feel, will fill an important gap in the advancement and applications of the information concept. At the same time, it is hoped that the proposed research will enable us to present a detailed and systematic development of information theory, its connexion with various mathematical theories, and its applications to certain fields of interest to applied workers.

Furthermore, it is hoped that it will be possible for us to develop methods for obtaining the amount of information contained in certain kinds of stochastic processes and probability spaces, and to present statistical techniques which would enable one to do so.

This, we believe, will prove to be of great value in various applied fields, apart from the significance of research for theoretical purposes themselves.

REFERENCES

(1) Shannon, C.E.

"A Mathematical Theory of Communication"
Bell System Technical Journal, Vol. 27 (1948).

(2) Shannon, C.E. and Weaver, W.

"The Mathematical Theory of Communication"
University of Illinois Press, 1949.

(3) Kullback, S.

"Information Theory and Statistics"
Wiley Publications, N.Y., 1959.

(4) Halmos, P.

"Entropy in Ergodic Theory"
Mimeographed notes, University of Chicago, 1959.

(5) Kinney, J.

"Singular Functions Associated with Markov Chains"
Proceedings of the American Mathematical Society, Vol. 9, 1958.

(6) Billingsley, P.

"Hausdorff Dimensions in Probability Theory," I & II
Illinois Journal of Mathematics, Vols. 4 & 5, 1960 and 1961.

(7) Chatterji, S.D.

"Certain Induced Measures and the Fractional Dimensions of their Supports"
To be published in Zeitschrift fur Wahrscheinlichkeitsrechnung. Available as technical report, Mathematics Research Center, U.S. Army, University of Wisconsin.

(8) Rényi, A.

"On the Dimension and Entropy of Probability Distributions"
Acta Mathematica, Academiae Scientiarum Hungaricae, Tomus X, 1959.

(9) Lorentz, G.G.

"Metric Entropy, Widths and Superpositions of Functions"
American Mathematical Monthly, Vol. 69, 1962.

(10) Khinchin, A.I.
Mathematical Foundations of Information Theory, Dover Publications, New York, 1957.

III.3 Resumés

Resumés of all the persons, except Mr. Bennie O. Flagg, working on the project have been included in the first semi-annual status report; Flagg's resumé is given below.

Name: Bennie O. Flagg, Jr.

Date and Place of Birth: July 26, 1940 Memphis, Tennessee

Marital Status: Single

Present Position: Research Assistant (NASA)

Educational Experience:

Degrees:

B.S. in Natural Science	LeMoyne College Memphis, Tennessee
M.A. (working on)	Kent State University Kent, Ohio

Teaching and Training:

Kent State University Kent, Ohio	Assistant Instructor 1963-1964

Fields of Present Interests: Algebra, Birth-and-death processes

Remarks: Member of American Mathematical Society